

Я. А. Ледовая, Р. В. Тихонов, В. Ю. Иванов,
Б. Р. Яминов, О. Н. Боголюбова

ОРГАНИЗАЦИОННО-МЕТОДИЧЕСКИЕ ВОПРОСЫ СБОРА ДАННЫХ В ОНЛАЙН-ИССЛЕДОВАНИИ ПОВЕДЕНИЯ ПОЛЬЗОВАТЕЛЕЙ СОЦИАЛЬНОЙ СЕТИ «ФЕЙСБУК» ИЗ РОССИИ И США*

Подробно описываются основные этапы и частично основные результаты проекта, в рамках которого с помощью специального онлайн-приложения проводился сбор психологических, демографических и текстовых данных пользователей социальной сети «Фейсбук» из России и США. Психологические характеристики респондентов оценивались через онлайн-опросники со встроенной автоматической обратной связью (в том числе Короткий опросник Темной триады, опросники выраженности отчуждения моральной ответственности, субъективного благополучия «ВОЗ-5», удовлетворенности жизнью SWLS Э. Динера, короткая шкала для скрининга посттравматических симптомов PC PTSD). Кроме того, с согласия участников приложение позволяло скачивать демографические данные из аккаунтов и тексты публичных постов пользователей, которые впоследствии анализировались лингвистическими методами. По нашим данным, впервые в России приведено подробное описание этапов подобного исследования, включая данные о затратах на рекламу приложения среди пользователей социальной сети «Фейсбук» в двух странах, и представлена часть содержательных результатов проекта, завершающегося в СПбГУ. Упомянуты необходимые для учета будущими исследователями важные нюансы организации такого подхода к сбору данных (в том числе временные, финансовые, связанные с созданием программного обеспечения и с формулировками вариантов обратной связи для респондентов), а также возможные технические, организационные и методические сложности и способы их преодоления с опорой на опыт собственной работы. Библиогр. 25 назв. Табл. 2. Ил. 4.

Ключевые слова: методы сбора данных, этапы онлайн-исследования, социальные сети, Фейсбук, поведение пользователей социальных сетей, сбор данных через приложение, обратная связь в интернет-исследованиях.

Ледовая Янина Александровна — Санкт-Петербургский государственный университет, Российская Федерация, 199034, Санкт-Петербург, Университетская наб, 7–9; y.ledovaya@spbu.ru

Тихонов Роман Вадимович — аспирант, Санкт-Петербургский государственный университет, Российская Федерация, 199034, Санкт-Петербург, Университетская наб, 7–9; roman.tikhonov@me.com

Иванов Виктор Юрьевич — Интернет-компания «Мэйл.ру», Российская Федерация, 125167, Москва, Ленинградский пр., 39; viktor.ivanov@corp.mail.ru

Яминов Булат Равилевич — XebiaLabs, Нидерланды, 1213VB, Хилверсум, Laapersveld 27; bulat.yaminov@gmail.com

Боголюбова Ольга Николаевна — кандидат психологических наук, Университет Кларксон, США, 13699, штат Нью-Йорк, Потсдам, а/я 5825; obogolyu@clarkson.edu

Ledovaya Yanina A. — Saint Petersburg State University, 7–9, Universitetskaya nab., St. Petersburg, 199034, Russian Federation; y.ledovaya@spbu.ru

Tikhonov Roman V. — postgraduate, Saint Petersburg State University, 7–9, Universitetskaya nab., St. Petersburg, 199034, Russian Federation; roman.tikhonov@me.com

Ivanov Viktor Yu. — Mail.ru Ltd., Russia, 125167, Moscow, Leningradsky pr., 39; viktor.ivanov@corp.mail.ru

Yaminov Bulat R. — XebiaLabs, 27, Laapersveld, Hilversum, 1213VB, The Netherlands; bulat.yaminov@gmail.com

Bogolyubova Olga N. — PhD, Clarkson University, Department of Psychology, P.O.Box 5825, Potsdam, NY, 13699, USA; obogolyu@clarkson.edu

* Работа выполнена при поддержке гранта СПбГУ 8.38.351.2015 «Стресс, здоровье и психологическое благополучие в социальных сетях: кросскультурное исследование».

© Санкт-Петербургский государственный университет, 2017

ORGANISATIONAL AND METHODOLOGICAL ISSUES OF DATA COLLECTION IN AN INTERNET BASED STUDY OF FACEBOOK USERS FROM RUSSIA AND USA

The article presents a detailed description of the major steps and a description of the basic findings of the project aimed to gather psychological, demographic and text data of Facebook users from Russia and the USA by means of a special web-application and subsequently analyze them. Psychological characteristics were evaluated with the help of online questionnaires with the built-in feedback (e. g. Short Dark Triad, moral disengagement, subjective wellbeing WHO-5, SWLS, PC PTSD scales). Besides, on participants' consent, the application downloaded demographic data from their accounts and their public wall posts' texts that were later analyzed with linguistic methods. The article for the first time in Russia presents the description of each step of that kind of research, including the ads campaign on Facebook expenses for Russia and USA (which allowed to distribute the link to the application). It also presents a number of significant research findings of the project being finalized at St. Petersburg State University. We mention the necessary points in data collection for the colleagues who plan to organize similar research, including those in terms of time, money, software development, feedback phrasing. We also draw attention to the possible technical, organizational and methodological complications and also to the means to overcome them based on our experience. Refs 25. Tables 2. Figs 4.

Keywords: methods of data collection, the milestones of an internet based study, social networks, Facebook, social networks users' behavior, data collection by means of application, the feedback in internet based studies.

Введение

Как было указано в предыдущей статье исследовательского коллектива [1], за последние 10–15 лет изменился образ жизни более чем половины населения мира: имея возможность регулярной коммуникации посредством социальных сетей, современные жители развитых стран не менее двух часов в день проводят за общением и получением информации через социальные сети («Фейсбук», «Инстаграм», «Твиттер», «ВКонтакте», «Одноклассники» и др.) [2; 3]. Молодые люди тратят на эти формы коммуникации еще больше времени [4], ведь для них они стали доступны в раннем возрасте, а значит, естественным образом встроились в образ жизни и каждодневный опыт. У этого нового когнитивно- и социально-психологического феномена есть важная техническая составляющая: история поведения пользователя, его «цифровые следы» (“digital footprints”). При наличии технических возможностей (и иногда в силу различных ограничений, введенных руководством различных социальных сетей) эти следы можно исследовать. Такой подход к сбору данных и в целом к созданию и проведению психологических и междисциплинарных исследований набирает популярность [1; 5; 6]. Эту информацию в своих интересах используют маркетинговые компании и службы разведки и контрразведки разных стран [7]. Но и исследователи, представители социальных наук и лингвисты, а также математики, программисты и специалисты по работе с данными начали изучать поведение людей в социальных сетях в попытках найти и проанализировать внешне наблюдаемые и зафиксированные следы пользовательской истории. Для осуществления подобных проектов не нужны затраты на помещения, оборудование, договоренности с испытуемыми: важно наличие компьютеров, сервера, хостинга для хранения данных и доступ в Интернет. Кроме того, для привлечения респондентов в социальной сети необходимо оплачивать рекламу объявления, содержащего ссылку на приложение, позволяющее принять участие в исследовании, но эти затраты

в целом невелики. Пожалуй, самый важный ресурс — это наличие в коллективе проекта специалистов из разных областей знания и коммуникация между ними.

В статье будут описаны особенности процесса сбора данных в социальной сети «Фейсбук» на основе опыта реализации проекта, проводящегося в СПбГУ. В этом проекте исследуются психологическое благополучие и ряд личностных особенностей пользователей из России и США в связи с их поведением и публичными текстами в социальной сети. Подобные онлайн-исследования становятся все более актуальными, так как одной из угроз психическому здоровью и психологическому благополучию пользователей социальных сетей является нежелательная агрессия, которой они могут подвергнуться в социальных сетях. Анонимность, возможность мгновенного обмена сообщениями, доступность коммуникации практически с любыми пользователями создают как возможности, так и угрозы. Россия занимает пятое место из 25 стран по показателям онлайн-буллинга, и 49% российских детей и подростков от 8 до 17 лет сообщают об опыте жертвы кибербуллинга (тогда как в среднем этот показатель составляет 37%), а 33% признают, что выступали в роли агрессора (по сравнению со средней долей 24% среди 25 стран) [8]. В литературе появилось понятие «кибербуллинг», или «троллинг», обозначающее целенаправленное (чаще вербальное) воздействие некоторых пользователей на эмоциональное состояние других партнеров по онлайн-коммуникации с целью вызвать негативные эмоции, тревогу, страх, огорчение [9]. И именно изучение возможных причин такого опасного для психического здоровья обычных пользователей поведения, а также задача поиска языковых коррелятов поведения как социально-сетевых агрессоров, так и тех, кто мог пострадать от их нападений, послужило мотивами для организации проекта.

Описываемое исследование, по нашим данным, является первым в своем роде в академической среде в России. В его основе — использование еще не применявшихся в отечественной психологической практике методов и методик. Поэтому основная цель статьи — описать этот достаточно новый исследовательский подход, упомянуть возможные сложности, которые могут встретиться на пути коллег, заинтересовавшихся подобным методом сбора данных в социальных сетях, а также зафиксировать пути их решения, предложенные в ходе работы участниками нашего научного проекта.

В рамках проекта данные собирались дважды: осенью 2015 г. в России и осенью-зимой 2016 г. в России и США. Организационно-методические вопросы сбора данных будут описаны в хронологическом порядке.

Используемые методики и вопросы

Основной новацией описываемого исследования можно назвать метод сбора данных в социальной сети «Фейсбук» при помощи специально созданной программы-приложения (она будет описана ниже). В 2015 г. в СПбГУ стартовал междисциплинарный проект, в рамках которого впервые на российской выборке с использованием метода сбора данных пользователей социальной сети «Фейсбук» через приложение исследовались «цифровые следы» респондентов (в частности, характеристики их публичных текстов, постов). Затем эти результаты соотносились с данными психологических опросников, собранными с помощью этого же приложения.

На следующем этапе англоязычные оригиналы указанных опросников использовались для аналогичного сбора данных среди пользователей сети «Фейсбук» из США, и в дальнейшем проводилось сопоставление результатов, полученных в двух странах.

Для сокращения времени заполнения опросников респондентами и для большей вероятности получить максимальное возможное количество завершенных сессий работы с приложением мы подбирали максимально короткие опросники. В исследование были заложены гипотезы о возможном влиянии ряда негативно окрашенных личностных особенностей («Темная триада», отчуждение моральной ответственности), а также присутствия в жизненном опыте травматичных событий и их последствий на агрессивное поведение в социальной сети. Для этого несколько еще не использовавшихся в России опросников прошли процедуру двойного перевода, первичного и повторного тестирования офлайн-респондентов и проверки внутренней согласованности их структуры.

Психологическое благополучие оценивалось с помощью шкалы субъективного благополучия ВОЗ-5 [10], переведенной ее разработчиками на русский язык [11]. При сборе данных в 2016 г. использовалась также шкала удовлетворенности жизнью Э. Динера [12] в переводе и адаптации коллектива исследования [13].

Для измерения *«темных» личностных черт* использовались две методики: 1) опросник «Темная Триада личностных черт» (нарциссизм, макиавеллизм, неклиническая психопатия) Джонса и Полхуса [14] в адаптации М. С. Егоровой и коллег [15]¹; 2) опросник выраженности отчуждения моральной ответственности (совокупности когнитивных механизмов, избирательно активирующих моральные запреты субъекта по отношению к себе при совершении социально осуждаемых, вредоносных поступков [16]) авторства С. Мур [17] в переводе и адаптации коллектива проекта и их коллег [18].

Нами использовалась короткая шкала для скрининга *посттравматических симптомов* Primary Care PTSD Screen [19] в переводе и адаптации коллектива исследования [13]. В шкале четыре вопроса о возможных последствиях травматичных событий, произошедших в жизни. Например: «В течение жизни переживали ли Вы какое-то событие, которое было настолько пугающим, ужасным или устрашающим, что за последние 30 дней у Вас были ночные кошмары, связанные с этим событием, или оно всплывало в памяти против вашей воли?» (с вариантами ответов «да» или «нет»).

Помимо стандартизированных опросников, использовались также отдельные вопросы о *демографических характеристиках* респондентов:

- вопрос о трудоустройстве с вариантами ответов («работаю как наемный работник», «работаю на себя», «домохозяйка», «студент», «пенсионер», «не работаю по другой причине»);
 - вопрос об уровне образования с вариантами ответов (неполное среднее, среднее, среднее профессиональное, неполное высшее, высшее, ученая степень).
- Также использовались вопросы для *самооценки поведения* пользователей:
- вопрос о частоте использования социальной сети (с вариантами «обычно не бываю в Фейсбуке», «бываю раз в месяц», «бываю раз в неделю»,

¹ В используемой нами версии мы оставили вопрос из оригинальной шкалы Джонса и Полхуса «Мне нравится секс с едва знакомыми людьми» вместо варианта русскоязычной версии М. С. Егоровой и коллег «Мне нравится подкалывать неудачников».

«бываю раз в день», «захожу несколько раз в день», «нахожусь в нем постоянно»);

- вопросы об опыте столкновения с интернет-агрессией в роли агрессора (3 вопроса), жертвы (3 вопроса) и свидетеля (1 вопрос), заданные косвенным образом, например: «Случалось ли так, что вы писали комментарии, содержащие насмешки, оскорбительные высказывания, угрозы к заметкам, фотографиям и другому контенту других людей в Фейсбуке?» с пятью вариантами ответа от «никогда» до «часто»;
- три вопроса о возможном опыте насилия (физическое нападение, домашнее насилие, сексуальное насилие), также с пятью вариантами ответа от «никогда» до «часто».

С согласия пользователей мы также получили со страницы профиля в социальной сети их возраст, пол, город (регион) проживания, подписки на публичные страницы (собирались в 2016 г.; не более 50 с каждого пользователя), публично открытые записи (посты) и скопированные записи (репосты).

Организация процедуры сбора данных

Для респондента взаимодействие с онлайн-приложением выглядело следующим образом (см. в Приложении скриншоты интерфейса опроса, иллюстрирующие приведенные ниже пункты 1, 3, 5).

1. Респондент видит в ленте новостей рекламное объявление, приглашающее принять участие в психологическом исследовании (ниже оно будет описано подробнее), кликает на него.
2. Приложение запрашивает доступ к ленте публикаций пользователя, дате рождения, месту жительства, его идентификатору.
3. Пользователю показывается стартовая страница приложения, на которой коротко объяснены цели исследования, приведена короткая версия информированного согласия (есть ссылка на отдельную страницу с его полным текстом), приведено обещание по окончании участия в опросе предоставить «персональный психологический портрет и индивидуальные рекомендации», дан адрес специально созданного почтового ящика проекта, на которой можно присылать вопросы (и, действительно, на несколько десятков вопросов респондентов впоследствии сотрудники проекта отвечали). Перейти дальше, к вопросам, можно только после принятия условий, поставив галочку в поле «Я принимаю условия соглашения».
4. Все данные, к которым пользователь дал доступ, отправляются на сервер приложения.
5. Пользователь начинает отвечать на вопросы приложения (это занимало примерно 10–15 минут, вопросов было 61 в 2015 г. и 66 в 2016 г.). По окончании выводилась страница с благодарностью за участие и обратной связью («психологическим портретом»). Формулировки и дизайн обратной связи были разными в 2015 и 2016 гг., об этом будет написано ниже.

Подготовка и три процедуры сбора данных (один в России в 2015 г., второй в России в 2016 г. и третий в 2016 г. в США) заняли примерно два года. Последовательность этапов реализации проекта представлена в таблице 1.

**Таблица 1. Основные этапы работы над подготовкой и сбором данных
(февраль 2015 — сентябрь 2017)**

Название и продолжительность этапа	Содержание этапа
1. Подготовительный (около восьми месяцев)	<ol style="list-style-type: none"> 1. Перевод, обратный перевод методик, проведение теста и ретеста на студенческой аудитории («Отчуждение моральной ответственности», «Темная триада» (версия не вошла в исследование), «Удовлетворенность жизнью», SWLS, «Шкала для скрининга посттравматических симптомов, PC-PTSD») и ряда других коротких шкал, отражающих опыт переживания травмирующих событий, которые не вошли в основное исследование. 2. Подготовка документов для Этического комитета, получение разрешения. 3. Регистрация в качестве разработчика на платформе «Фейсбук», создание приложения, получение одобрения приложения от программистов и юристов «Фейсбук», подготовка формулировок обратной связи
2. Первый сбор данных (около двух недель)	<ol style="list-style-type: none"> 1. Отладка работы приложения с включенными в него опросниками и автоматической обратной связью. 2. Создание страницы приложения в сети «Фейсбук», регистрация в качестве рекламодателя в сети «Фейсбук», создание макета рекламного объявления и короткого рекламного текста к нему. 3. Подключение банковской карты к рекламному аккаунту, настройка таргетирования показов рекламного объявления и лимитов ежедневных трат. 4. Запуск рекламной кампании, контроль набора данных и сопутствующих трат
3. Обработка и описание результатов (около шести месяцев)	<ol style="list-style-type: none"> 1. Проверка качества полученных ответов и предобработка данных. 2. Статистический анализ психологических данных и описание полученных результатов. Выявлены взаимосвязи опыта переживания насилия с психологическим благополучием и психологическим дистрессом [20]. Обнаружены взаимосвязи личностных черт с самоотчетами об агрессивном поведении в социальной сети [21]. 3. Применение методов математической лингвистики для автоматического подсчета морфологических и лексических характеристик текстов; кластеризация слов на основе семантических признаков. Были выявлены специфические лексические, морфологические и семантические предикторы агрессивного онлайн-поведения [21; 22]
4. Оптимизация приложения (около двух месяцев)	<ol style="list-style-type: none"> 1. Оптимизация опросника для контроля качества ответов респондентов: добавление вопросов, диагностирующих социальную желательность (3), и вопроса-ловушки (1), добавление модуля, измеряющего время предъявления вопросов. 2. Модификация формы обратной связи: программирование формулы для расчета, добавление рисунков мифологических персонажей и описаний типажей, возможность поделиться своими результатами, расширяя таким образом приложение. 3. Добавление опросника SWLS Э. Динера. 4. Увеличение количества собираемых текстовых данных
5. Второй сбор данных (один месяц)	<ol style="list-style-type: none"> 1. Запуск рекламной кампании в РФ, таргетированной на разные города и в большей степени на более молодую аудиторию мужского пола. 2. По окончании сбора данных в РФ — размещение в приложении вопросов и обратной связи на английском языке

Название и продолжительность этапа	Содержание этапа
6. Третий сбор данных — в США (около двух недель, плюс перевод обратной связи перед этим — около месяца)	1. Перевод страницы, инструкции, вопросов и обратной связи на английский язык. 2. Таргетирование показа объявлений в разных штатах США, а также таргетирование на более молодую аудиторию
7. Обработка и описание результатов (около десяти месяцев)	1. Препроцессинг и исключение результатов, не прошедших процедуру оценки качества ответов. 2. Статистический анализ данных и описание полученных результатов. 3. Применение методов машинного обучения для предсказания личностных черт, благополучия и особенностей поведения пользователей на основе их публичных текстов со страницы социальной сети [23; 24]

Опишем эти этапы и их содержание более подробно.

Подготовительный этап был связан с отбором и русифицированием психологических методик, получением разрешения от Этического комитета, а также с созданием программы-приложения и регистрацией ее в качестве приложения, одобренного юристами и техническими специалистами сети «Фейсбук»².

К *этапу первого сбора данных*, после создания и отладки приложения со встроенными опросниками и автоматической краткой обратной связью, можно отнести запуск приложения и рекламной кампании в социальной сети «Фейсбук» (с 8 по 20 ноября 2015 г.). Он включал в себя несколько шагов, связанных с: а) созданием специальной страницы приложения в социальной сети (<https://www.facebook.com/wellbeing.and.networks/>), на которой коротко описывались цели исследования и присутствовало рекламное объявление, призывающее принять участие в исследовании, а также ссылка на собирающее данные приложение (<https://apps.facebook.com/psytest>); б) проектированием дизайна и короткого слогана рекламного объявления (см. прил. 1); в) регистрацией аккаунта «рекламодателя» в сети «Фейсбук»; г) настройкой и таргетированием показов этого объявления определенным аудиториям (в исследовании 2015 г. — жители России старше 18 лет, в 2016 г. таргетирование было более специализированным); д) подключением банковской карты с достаточной суммой на счете к аккаунту («Фейсбук» принимает оплату только через банковскую карту, причем деньги перечисляются в Ирландию, необходима банковская карта с такой возможностью) и контролем оплаты показов (в 2015 г. было потрачено на рекламу 76 000 руб.); е) отслеживанием эффективности набора респондентов (см. табл. 2).

² Отдельно остановимся на подготовке формулировок обратной связи. Пользователям возвращалось короткое описание их самочувствия по результатам шкалы ВОЗ-5 (с характеристиками отличного, умеренно-положительного и неблагоприятного характера). Для отнесения результатов респондентов к одному из трех описанных уровней программой высчитывалась сумма по всем вопросам шкалы, которая сопоставлялась с найденными в литературе данными по достаточно большой выборке жителей стран Скандинавии [27]. Также пользователи получали обратную связь по итогам выраженности трех субшкал шкалы «Темная триада»: в первый год были взяты средние данные из статьи М. С. Егоровой и коллег о сборе данных на российской офлайн-выборке, во второй год были изменены дизайн и расчет обратной связи по этой шкале.

Таблица 2. Показатели эффективности трех рекламных кампаний, использованных для сбора данных в сети «Фейсбук»

Показатель	Этап сбора данных		
	Первый (2015, РФ)	Второй (2016, РФ)	Третий (2016, США)
Пользователи, которым показывалась реклама	191 911	105 372	214 784
Перешли в приложение по рекламе	17 930	12 436	2757
Согласились участвовать в исследовании	8396	8367	1316
Полностью завершили опрос	6724	3436	645
Вошли в итоговую выборку	6724*	3341	629
Стоимость данных одного участника (руб.)	11,3	8,3	155

* В 2015 г. участники с социально желательными или некачественными ответами не исключались из анализа, так как вопрос-ловушка и шкала социальной желательности были добавлены позднее.

На этапе *обработки и описания полученных данных* мы увидели возможность улучшить процедуру сбора данных и качество выборки. Дело в том, что распределение по полу оказалось неравномерным — 77,9 % выборки составили участники женского пола, а средний возраст выборки составил 45 лет ($\pm 11,58$). Данные текстов постов по ряду технических причин были получены только от 1972 участников (в среднем у этих пользователей было скачано по 7,67 постов). После этого мы начали работать над оптимизацией процедуры сбора данных.

В процессе *оптимизации приложения* были удалены несколько вопросов о негативных переживаниях и их последствиях (чтобы пользователи были более мотивированы отвечать на вопросы и заполнить опросник до конца), добавлены три вопроса, отслеживающих социальную желательность, один вопрос-ловушка, отсекающий тех, кто вообще не читает вопросы и отвечает наобум. Кроме того, получив некоторое количество комментариев участников под рекламным объявлением о том, что обратная связь по двум методикам (опросник «Темная триада» и шкала субъективного благополучия ВОЗ-5, для обратной связи брались результаты только по ним) иногда содержала противоречивые утверждения, мы изменили форму подачи обратной связи.

Также в процессе оптимизации приложения для подготовки сбора данных во второй раз была создана более яркая и упрощенная форма обратной связи. Сделано это было в том числе для того, чтобы не усложнять восприятие информации на темы, связанные с «темными» чертами самих пользователей, а также для того, чтобы после нескольких десятков вопросов на достаточно чувствительные темы они получили описание, не усугубляющее возможный дистресс. Среди трех субшкал опросника «Темная триада» программой выбиралась шкала, по которой респондент набрал больше всего баллов (с поправкой в виде числового коэффициента на то, что в российской выборке 2015 г. показатели шкалы макиавеллизма в среднем

были самыми высокими, а по шкале психопатии — самыми низкими), и тогда респонденту выводился на экран «портрет» одного из трех типажей с достаточно шутливым описанием. Были найдены аналоги носителей трех ярко выраженных «темных черт» — персонажи древнегреческих мифов. Кроме того, для мужчин и женщин были подобраны боги и герои разных полов, а в Интернете были найдены черно-белые изображения этих персонажей. Таким образом, для женщин обратная связь могла быть о том, что «по своим личностным чертам Вы больше всего похожи на Афродиту (Геру, Медузу Горгону)...», а для мужчин — о том, что «по своим личностным чертам Вы больше всего похожи на Нарцисса (Одиссея, Минотавра)». Далее следовал абзац, описывающий типичное поведение людей с достаточно яркой выраженностью нарциссизма (макиавеллизма, неклинической психопатии), причем акцент делался на сильных, положительных сторонах их поведения. Также в новой версии содержалась возможность поделиться своими результатами с друзьями в социальной сети, и эта опция помогла нам привлечь довольно большое число участников исследования бесплатно (примерно треть от общего количества). Упрощение и сведение результатов к описанию типа были выбраны осознанно. Отсылка к мифологическим персонажам вызывает положительные эмоции из-за факта узнавания уже известного, а яркий типаж упрощает запоминание типичных паттернов приписанного нами им поведения. (Иногда когнитивные искажения работают во благо.)

Во время *второго сбора данных* в России (месяц, с 8 октября по 9 ноября 2016 г.) было уделено внимание таргетированию показа объявлений на пользователей мужского пола, а также на более молодых пользователей. Изначально при одинаковом таргетинге на всех пользователей из России старше 18 лет женщины и люди старше 40–45 лет переходили в приложение по объявлению чаще. В связи с этим через некоторое время рекламная кампания была перезапущена с таргетингом только на мужчин и на людей до 45 лет (это позволяют возможности аккаунта рекламодателя в сети «Фейсбук»). Кроме того, таргетинг на некоторое время направлялся на крупные города (Москва, Санкт-Петербург, города-миллионники), откуда, по данным панели рекламного аккаунта, не было респондентов. Мы объяснили это тем, что коммерческие рекламодатели таргетируют свою рекламу на эти города, и их жители в большей степени загружены разнообразными рекламными объявлениями в ленте новостей сети «Фейсбук», а реклама приложений автоматически направляется на максимально доступную и экономически выгодную аудиторию.

Охват пользователей оказался менее масштабным, чем в 2015 г. (см. табл. 2), однако эффективность рекламы повысилась примерно на 26 %. Кроме того, возможность поделиться результатами исследования с друзьями увеличила охват пользователей без дополнительных затрат. Качество полученных данных также удалось повысить за счет того, что в результирующую выборку попали только пользователи, которые прошли предварительный отбор (учитывалась социальная желательность, скорость ответов и внимательность чтения формулировок вопросов). Затраты на рекламу составили примерно 28 000 руб., почти в три раза меньше, чем при первом сборе данных.

При *сборе данных среди пользователей из США* с 6 по 16 декабря 2016 г. (для этого были переведены тексты начальной страницы с кратким информированным согласием и тексты вариантов обратной связи, а все опросники изначально были

англоязычными) была затрачена существенная сумма, но при этом получены более количественно скромные результаты, чем при сборе данных в российском сегменте сети «Фейсбук» (см. табл. 2). На показы рекламных объявлений было потрачено 97 500 руб., а средняя стоимость данных, полученных в США, составила 155 руб. за каждого участника (или 2,5 долл. США) (что в целом не очень дорого по сравнению с вариантами стоимости сбора аналогичных данных, описанными в статье Косински и соавторов (от 1,5 до 33 долл., в среднем — 13,75 долл.) [5]). Но это почти в 20 раз выше стоимости сбора данных в российском сегменте сети «Фейсбук». Возможно, это произошло из-за более высокой активности рекламодателей в США, чем в России. Кроме того, упоминание российских психологов как организаторов исследования могло снизить привлекательность рекламы.

Для того чтобы выборка была более уравновешенной по полу и возрасту, необходимо было таргетировать показ объявления на аудиторию 25–34 лет и на женщин (тогда как в России пользователей-женщин, наоборот, было больше).

Описав последовательность сбора данных, осуществляемую и контролируруемую психологами проекта, выделим и опишем два важных, на наш взгляд, технически значимых этапа в работе: этап подготовки приложения в начале (осуществляется веб-программистом в постоянной коммуникации с психологами и лингвистами) и этап препроцессинга данных перед проверкой гипотез по итогам их сбора (осуществляется психологом, специализирующимся на обработке данных).

Краткое описание особенностей программы-приложения для сбора данных

Перейдем к обобщенному описанию специального приложения, которое осуществляло онлайн-тестирование, выдавало пользователям краткую обратную связь о них и с их разрешения собирало тексты их открытых постов, а также минимальную демографическую информацию из профиля и «лайки страниц» — подписки на публичные страницы компаний, организаций, медиа-персон в сети «Фейсбук».

Приложение, с помощью которого собирались данные, в нашем случае было создано для веб-версии «Фейсбук» (т. е. для тестирования на компьютере или ноутбуке, а не через браузеры смартфонов или планшетов). Социальная сеть через свой API-интерфейс (программный интерфейс социальной сети) дает возможность приложению запросить доступ к ленте публикаций пользователя, его месту жительства, информации об интересных ему страницах, личной информации — дате рождения, семейном положении и т. д. «Фейсбук» очень серьезно подходит к вопросам безопасности личных данных пользователя, поэтому для получения этих данных с помощью приложения нужно двойное подтверждение — вначале, на этапе регистрации приложения в сети, для каждого конкретного пункта собираемых через профиль данных требуется разрешение самой социальной сети (ее программистов и юристов), причем необходимо обосновать причину, по которой эти данные нужны разработчику, и причину полезности именно для пользователя факта передачи этой информации нам (обоснования, написанные на английском языке, отправляются из аккаунта программиста-разработчика в специальном разделе социальной сети, и в нашем случае пришлось сделать несколько итераций, прежде чем было получено разрешение). Вторым этапом является разрешение на скачивание этих

данных от самого пользователя в момент первого перехода в приложение. Кроме того, в сети «Фейсбук» некоторые данные так и остались за пределами API и не могут быть получены никаким официальным образом — это список друзей пользователя (можно получить лишь список друзей, которые также переходили в это приложение), его личные сообщения, информация о его действиях в «Фейсбук».

Еще одна составляющая безопасности приложения по требованиям сети «Фейсбук» — это наличие и ежегодное продление SSL-сертификата для передачи зашифрованных данных по протоколу HTTPS. О его покупке и ежегодном обновлении нужно тоже помнить. Кроме того, необходим хостинг и сервер, где будут храниться данные. Под данные нашего проекта (мы собирали данные дважды в России и один раз в США) зарезервировано 4 Гб.

В качестве системы управления базами данных используется MySQL, для которой у использованного в проекте языка PHP есть много встроенных функций. MySQL является основным хранилищем как для ответов пользователей, так и для данных, полученных через API Facebook.

Важно отметить, что подобные исследования со стационарных компьютеров через веб-интерфейс постепенно будет все сложнее проводить. Пользователи все чаще переходят на использование смартфонов, а заполнять опросник около 15–20 минут и не отвлекаться от мобильного устройства — паттерн, противоречащий самой идее мобильности. В связи с этим для проведения интернет-исследований мы рекомендуем максимально сокращать количество вопросов, программировать возможность прервать процедуру и вернуться к ней, а также по возможности геймифицировать взаимодействие с приложением, чтобы удерживать внимание пользователей (как минимум добавить строку прогресса). А разрабатывая веб-приложения, быть готовыми к тому, что относительно просто будет собрать лишь данные пользователей старшего поколения, которые чаще пользуются стационарными компьютерами.

Препроцессинг полученных данных

В качестве примера приведем применяемую нами процедуру предобработки данных, полученных в 2016 г. среди пользователей русскоязычного сегмента сети «Фейсбук». Она включала в себя следующие этапы:

- 1) исключение неполных ответов;
- 2) исключение ответов невнимательных респондентов с помощью вопроса-ловушки;
- 3) исключение ответов пользователей с высокой социальной желательностью;
- 4) исключение пользователей, которые отвечали на вопросы быстрее обычного.

Из 9683 человек, согласившихся принять участие в исследовании в 2016 г., только 4240 человек (44 %) смогли завершить опрос, ответив на все 66 вопросов. Интересно, что более половины из всех отказавшихся (61 %) не стали отвечать даже на первый вопрос, а еще около трети из всех отказов (34 %) произошли при ответе на первые 20 вопросов. И только 5 % из всех отказавшихся прекратили участие начиная с 21-го и заканчивая 65-м вопросом.

Участники достаточно внимательно подходили к чтению формулировок вопросов. Только 2,5 % участников, завершивших опрос, неверно ответили на

вопрос-ловушку, в котором требовалось выбрать обозначенный в формулировке вопроса вариант ответа («В этом вопросе выберите вариант “часто”»). Ответы этих респондентов были исключены из дальнейшего анализа.

Шкала социальной желательности включала три вопроса из опросника Кроуна-Марлоу [25] в адаптации Ю. Л. Ханина (использовалась пятибалльная шкала Лайкерта вместо оригинальной шкалы «да»/«нет»). Средний балл по модифицированной шкале оказался равен 8,2 ($\pm 2,4$). Мы исключили ответы участников с высокой социальной желательностью, т. е. тех, кто набрал 14 и 15 баллов из 15 возможных.

Мы также фиксировали время от момента загрузки каждого вопроса до выбора варианта ответа. Участников с большим количеством быстрых ответов (менее 5 сек.) мы исключили из анализа. В качестве критерия отсева мы использовали количество быстрых ответов, превышающее три стандартных отклонения от среднего по выборке.

Выводы

Мы описали последовательность подготовки и сбора психологических (а также частично текстовых) данных при помощи специального приложения для социальной сети «Фейсбук». Ниже приведем список ограничений процедур, которые мы провели, а также сформулируем полезные советы для тех исследователей, которые захотели бы воспользоваться нашим опытом и начать исследования, собирая «цифровые следы» пользователей социальных сетей.

Ряд ограничений, которые присутствовали в описанном исследовании.

Собирая данные в первый год, не имея подобного опыта, а также доступа к описанным кем-то пошаговым инструкциям о проведении такого рода проектов, наша группа неизбежно столкнулась с рядом технических недочетов, сказавшихся на качестве собранных в первый год исследования данных. Это небольшое количество текстов постов, которые получилось собрать, неравновесная по полу выборка без специального таргетирования рекламы, противоречия в обратной связи, вызвавшие вопросы респондентов.

В социальной сети «Фейсбук» в России достаточно сложно собрать данные молодых пользователей, потому что они чаще пользуются социальной сетью «ВКонтакте», и это сказалось на том, что средний возраст участников этого исследования — 45–46 лет.

Поскольку сеть «Фейсбук» на третьем месте по популярности в России (ею пользуются более 14 млн человек [26]), собранные в ней данные нельзя считать репрезентативно отражающими особенности всей популяции (но этот недостаток есть почти у каждой социальной сети).

Судя по загруженным в базу данных историям публикаций респондентов, участники исследования — это люди, регулярно откликающиеся на рекламу принять участие в различных онлайн-тестированиях (как правило, несерьезных, когда по особенностям профиля, текстов или частоте переписки с кем-то из друзей возвращается обратная связь о том, «какой ты герой сериала “Игра престолов”» и т. п.). Эта особенность данных также свидетельствует об их недостаточной репрезентативности и специфичности выборки (хотя позволяет проводить лингвистический анализ и сопоставление лингвистических и психологических данных).

Нам не удалось собрать достаточное количество данных пользователей из США для прямого сопоставления их с данными пользователей из России. Для этого необходимо либо тратить слишком существенные средства, либо менять рекламное объявление (при этом важно, чтобы оно сохраняло преемственность с тем, которое сработало на российской выборке).

Опираясь на свой опыт, мы бы хотели заострить внимание на некоторых нюансах проведения исследований в социальной сети.

Обратную связь (желательно в достаточно шуточной форме, потому что к этому привыкли пользователи социальных сетей) необходимо продумывать заранее и уделить ей время. Несложно сформулированная, необязательно исчерпывающая и в целом отражающая вопросы опросников обратная связь («психологический портрет») минимизирует неискренность в ответах и может стать стимулом для добровольного распространения респондентами ссылки на опрос.

Список вопросов должен быть недлинным, в пределах примерно 60–70, чтобы пользователи не уставали и не отказывались отвечать на них.

Для привлечения аудитории к серьезному академическому исследованию может сработать достаточно скромное рекламное объявление.

Важно иметь в виду, что все больше пользователей выходят в социальные сети через смартфоны, а это значит, что сессии ответов на вопросы должны быть как можно короче, желательно запрограммировать возможность прервать сессию и вернуться к ней позже. Также не лишними будут элементы геймификации в приложении, как минимум строка, отображающая количество отвеченных и оставшихся вопросов.

Важны вопросы на социальную желательность, автоматические ответы и фиксация времени ответа на один вопрос. Они помогут отсеять нежелательные результаты.

Также важно делать предварительную обработку и отсев данных с учетом описанных нами параметров.

Необходимо не забывать о затратах на создание и обновление SSL-сертификата для протокола HTTPS и хостинг.

В конце статьи важно написать о том, какие (судя по анализу литературных источников и отчасти с опорой на наш опыт работы в проекте) существуют ограничения и преимущества в подходе к описываемому варианту сбора данных в социальных сетях.

Ограничения процедуры сбора данных в социальных сетях.

1. Не все социальные сети (в частности, «Фейсбук») позволяют скачивать и анализировать связи между пользователями («дружбы»), а это важный и потенциально продуктивный элемент исследований такого характера.
2. Технически достаточно сложно проводить подобные исследования в сети «Фейсбук» из-за длительных согласований с администрацией сети и особенностей работы с API (технически проще работать с API в «Твиттере» и «ВКонтакте», но в России «Твиттер» нерепрезентативен, у него немного пользователей, а «ВКонтакте» недостаточно представлены люди старше 40 лет).

3. Чаще всего в таком исследовательском контексте доступны корреляционные модели исследований, без каузальных объяснений.
4. Собирать данные за границей (по нашему единичному опыту) довольно дорого по сравнению со стоимостью данных из России.
5. Не всегда в загруженных текстах (не репостах) пользователей представлены именно их тексты (они могут цитировать писателей, политиков и т. д.).

В заключение исходя из своего опыта работы опишем достоинства сбора данных в социальной сети.

Как показал наш опыт, собрать психологические, поведенческие (относительно поведения в социальной сети) и языковые данные на выборках, исчисляемых тысячами человек, в социальной сети «Фейсбук» — достаточно недорого и достаточно быстро. Это требует определенной длительной и аккуратной подготовки, но затем приносит свои продуктивные плоды. Вознаграждение респондентов, которые участвуют в исследованиях, проводимых отечественными психологами офлайн, составляет, по нашим сведениям, от 100 до 1000 руб. А собирая данные через рекламирование ссылки на приложение в сети «Фейсбук», мы потратили около 10 руб. на одного российского респондента и 155 руб. на одного респондента из США.

Для сбора данных, а потом для их обработки необходимо участие представителей других наук и специальностей — программистов, лингвистов, аналитиков данных. Такие проекты заведомо междисциплинальны. Но основные задачи в них ставят психологи.

Такой формат работы достаточно удобен (оперативен и относительно недорог) для сбора данных в нескольких странах, несмотря на то что попытка собрать данные в США в описываемом нами случае не увенчалась однозначным успехом.

Увеличивается количество пользователей русскоязычного сегмента сети «Фейсбук». Возможно, со временем выборки пользователей этой сети будут достаточно репрезентативными.

Как показывают данные наших и не только наших публикаций, тексты постов пользователей социальных сетей действительно отражают их самые разные личностные особенности (от факторов «Большой пятерки» до черт «Темной триады» и уровня психологического благополучия). Поэтому подобные исследования имеют немалую теоретическую и прикладную ценность.

На основе получаемых данных можно строить предсказательные модели, установив корреляции лексических, морфологических, синтаксических характеристик текстов постов с психологическими чертами их авторов. И такие прикладные модели могут стать основой для моделей теоретических в дальнейшем.

Собранные таким образом за одну-две рекламные кампании в социальной сети данные позволяют достаточно долго и продуктивно их обрабатывать и описывать силами специалистов из нескольких областей науки. Можно сказать, что материал для проверки немалого количества гипотез в течение еще нескольких лет в базах данных нашего проекта имеется.

В заключение хотим выразить признательность нашим коллегам — участникам проекта, помогавшим на разных этапах его реализации и оказавшим помощь при подготовке текста статьи: П. В. Паничевой, А. К. Москвичеву, А. Г. Чуриловой, М. А. Недошивиной, М. В. Салитовой, А. Б. Спектру.

1. Ледовая Я. А., Тихонов Р. В., Боголюбова О. Н. Социальные сети как новая среда для междисциплинарных исследований поведения человека // Вестн. С.-Петерб. ун-та. Психология и педагогика. 2017. Т. 7, № 3. С. 193–200.
2. Исследовательский холдинг Ромир. Социально-сетевая жизнь. 2015. URL: http://romir.ru/studies/670_1432155600/ (дата обращения: 21.08.2017).
3. Mander J. Daily time spent on social networks rises to over 2 hours // GlobalWebIndex. 2017. URL: <http://blog.globalwebindex.net/chart-of-the-day/daily-time-spent-on-social-networks/> (дата обращения: 12.08.2017).
4. Mander J., McGrath F. GWI Social Summary: GlobalWebIndex's quarterly report on the latest trends in social networking [Flagship Report Q1 2017]. URL: <http://www.globalwebindex.net/> (дата обращения: 30.08.2017).
5. Kosinski M., Matz S., Gosling S., Popov V., Stillwell D. Facebook as a research tool for the social sciences: Opportunities, challenges, ethical considerations, and practical guidelines // Am. Psychol. 2015. Vol. 70. No. 6. P. 543–556.
6. Guntuku S. C., Yaden D. B., Kern M. L., Ungar L. H., Eichstaedt J. C. Detecting depression and mental illness on social media: an integrative review // Curr. Opin. Behav. Sci. 2017. Vol. 18. P. 43–49.
7. Хачуян А. Можно ли определить террориста по его цифровому следу [видеозапись лекции] // YouTube. 2 июня. 2017. URL: <https://www.youtube.com/watch?v=vNmANAFB0EQ> (дата обращения: 29.10.2017).
8. Online Bullying Among Youth 8–17 Years Old — Russia [Report Q2-CY2012] / Cross-Tab Marketing Services & Telecommunications Research Group for Microsoft Corporation. 2012. URL: <https://www.microsoft.com/en-us/download/details.aspx?id=30148> (дата обращения: 10.06.2017).
9. Ybarra M. L., Mitchell K. J., Wolak J., Finkelhor D. Examining Characteristics and Associated Distress Related to Internet Harassment: Findings From the Second Youth Internet Safety Survey // Pediatrics. 2006. Vol. 118, N 4. P. e1169–e1177.
10. Bech P. Health-related quality of life measurements in the assessment of pain clinic results // Acta Anaesthesiol. Scand. 1999. Vol. 43, N 9. P. 893–896.
11. Индекс общего (хорошего) самочувствия ВОЗ / Psychiatric Research Unit. WHO Collaborating Center for Mental Health, Frederiksberg General Hospital. 1999. URL: https://www.psykiatri-regionh.dk/who-5/Documents/WHO5_Russian.pdf (дата обращения: 15.08.2017).
12. Diener E., Emmons R. A., Larsen R. J., Griffin S. The Satisfaction With Life Scale // J. Pers. Assess. Lawrence Erlbaum Associates, Inc. 1985. Vol. 49, N 1. P. 71–75.
13. Ледовая Я. А., Боголюбова О. Н., Тихонов Р. В. Стресс, благополучие и Темная триада // Психологические исследования. 2015. Т. 8, № 43. С. 5. URL: <http://psystudy.ru/index.php/num/2015v8n43/1185-ledovaya43.html> (дата обращения: 01.09.2017).
14. Jones D. N., Paulhus D. L. Introducing the Short Dark Triad (SD3) // Assessment. 2014. Vol. 21, N 1. P. 28–41.
15. Егорова М. С., Ситникова М. А., Паршикова О. В. Адаптация короткого опросника «Темной триады» // Психологические исследования. 2015. Т. 8, № 43. С. 1. URL: <http://psystudy.ru/num/2015v8n43/1181-egorova43.html> (дата обращения: 01.09.2017).
16. Bandura A. Social foundations of thought and action: A social cognitive theory. Englewood Cliffs, NJ: Prentice-Hall Inc, 1985. 544 p.
17. Moore C., Detert J. R., Klebe Treviño L., Baker V. L., Mayer D. M. Why employees do bad things: Moral disengagement and unethical organizational behavior // Pers. Psychol. 2012. Vol. 65, N 1. P. 1–48.
18. Ледовая Я. А., Тихонов Р. В., Боголюбова О. Н., Казенная Е. В., Сорокина Ю. Л. Отчуждение моральной ответственности: психологический конструктор и методы его измерения // Вестн. С.-Петерб. ун-та. Сер. 16. Психология. Педагогика. 2016. № 4. С. 23–39.
19. Prins A., Ouimette P., Kimerling R., Camerond R. P., Hugelshofer D. S., Shaw-Hegwer J., Sheikh J. I. The primary care PTSD screen (PC-PTSD): development and operating characteristics // Prim. Care Psychiatry. 2004. Vol. 9, N 1. P. 9–14.
20. Bogolyubova O., Tikhonov R., Ivanov V., Panicheva P., Ledovaya Ya. Violence Exposure, Posttraumatic Stress, and Subjective Well-Being in a Sample of Russian Adults // J. Interpers. Violence. 2017. P. 1–16. URL: <https://doi.org/10.1177/0886260517698279> (дата обращения: 01.10.2017).
21. Bogolyubova O., Panicheva P., Tikhonov R., Ivanov V., Ledovaya Ya. Dark Personalities on Facebook: Harmful Online Behaviors and Language // Comput. Human Behav. Elsevier Ltd, 2018. Vol. 78. P. 151–159. URL: <https://doi.org/10.1016/j.chb.2017.09.032> (дата обращения: 01.10.2017).

22. Panicheva P., Ledovaya Ya., Bogolyubova O. Lexical, Morphological and Semantic Correlates of the Dark Triad Personality Traits in Russian Facebook Texts // Proc. Artif. Intell. Nat. Lang. Conf. (AINL 2016). 2016. P. 1–8. URL: <http://ieeexplore.ieee.org/document/7891860/> (дата обращения: 01.10.2017).
23. Panicheva P., Mirzagitova A., Ledovaya Y. Semantic Feature Aggregation for Gender Identification in Russian Facebook // AINL 2017: Artificial Intelligence and Natural Language / eds A. Filchenkov, L. Pivovarova, J. Žizka. Communications in Computer and Information Science, vol 789. Springer, Cham. 2018. P. 3–15. URL: https://link.springer.com/chapter/10.1007/978-3-319-71746-3_1 (дата обращения: 01.12.2017).
24. Moskvichev A., Dubova M., Menshov S., Filchenkov A. Using Linguistic Activity in Social Networks to Predict and Interpret Dark Psychological Traits // AINL 2017: Artificial Intelligence and Natural Language / eds A. Filchenkov, L. Pivovarova, J. Žizka. Communications in Computer and Information Science. Vol. 789. Springer, Cham. 2018. P. 16–26. URL: https://link.springer.com/chapter/10.1007/978-3-319-71746-3_2 (дата обращения: 01.12.2017).
25. Crowne D. P., Marlowe D. A new scale of social desirability independent of psychopathology // J. Consult. Psychol. 1960. Vol. 24, N 4. P. 349–354.
26. Количество активных пользователей Facebook в России составляет 14,4 млн человек // Rambler News Service. 2016. URL: <https://rns.online/internet/Kolichestvo-aktivnih-polzovatelei-Facebook-v-Rossii-sostavlyayet-144-mln-chelovek-2016-05-13/> (дата обращения: 01.09.2017).
27. Guðmundsdóttir H., Ólason D., Guðmundsdóttir D., Sigurðsson J. F. A psychometric evaluation of the Icelandic version of the WHO-5 // Scand. J. Psychol. 2014. Vol. 55, N 6. P. 567–572.

Для цитирования: Ледовая Я. А., Тихонов Р. В., Иванов В. Ю., Яминов Б. Р., Боголюбова О. Н. Организационно-методические вопросы сбора данных в онлайн-исследовании поведения пользователей социальной сети «Фейсбук» из России и США // Вестник СПбГУ. Психология и педагогика. 2017. Т. 7. Вып. 4. С. 308–327. <https://doi.org/10.21638/11701/spbu16.2017.402>

References

1. Ledovaia Ia. A., Tikhonov R. V., Bogoliubova O. N. Sotsial'nye seti kak novaia sreda dlia mezhdistsiplinarnykh issledovaniy povedeniia cheloveka [Social Networks as a New Environment for Interdisciplinary Studies of Human Behavior]. *Vestnik SPbSU. Psychology and Education*, 2017, vol. 7, no. 3, pp. 193–200. (In Russian)
2. Issledovatel'skii kholding Romir. Sotsial'no-setevaia zhizn' [Research holding Romir. Life in social networks], 2015. Available at: http://romir.ru/studies/670_1432155600/ (accessed: 21.08.2017).
3. Mander J. Daily time spent on social networks rises to over 2 hours. *GlobalWebIndex*, 2017. Available at: <http://blog.globalwebindex.net/chart-of-the-day/daily-time-spent-on-social-networks/> (accessed: 12.08.2017).
4. Mander J., McGrath F. *GWJ Social Summary: GlobalWebIndex's quarterly report on the latest trends in social networking [Flagship Report Q1 2017]*. Available at: <http://www.globalwebindex.net/> (accessed: 30.08.2017).
5. Kosinski M., Matz S., Gosling S., Popov V., Stillwell D. Facebook as a research tool for the social sciences: Opportunities, challenges, ethical considerations, and practical guidelines. *Am. Psychol.* 2015, vol. 70, no. 6, pp. 543–556.
6. Guntuku S. C., Yaden D. B., Kern M. L., Ungar L. H., Eichstaedt J. C. Detecting depression and mental illness on social media: an integrative review. *Curr. Opin. Behav. Sci.*, 2017, vol. 18, pp. 43–49.
7. Khachuian A. Mozhno li opredelit' terrorista po ego tsifrovomu sledu [videozapis' lektsii] [Is it possible to identify a terrorist using his digital footprint]. *YouTube*, June, 2. 2017. Available at: <https://www.youtube.com/watch?v=vNmANAFB0EQ> (accessed: 29.10.2017). (In Russian)
8. *Online Bullying Among Youth 8–17 Years Old — Russia [Report Q2-CY2012] / Cross-Tab Marketing Services & Telecommunications Research Group for Microsoft Corporation*, 2012. Available at: <https://www.microsoft.com/en-us/download/details.aspx?id=30148> (accessed: 10.06.2017).
9. Ybarra M. L., Mitchell K. J., Wolak J., Finkelhor D. Examining Characteristics and Associated Distress Related to Internet Harassment: Findings From the Second Youth Internet Safety Survey. *Pediatrics*, 2006, vol. 118, no. 4, pp. e1169–e1177.
10. Bech P. Health-related quality of life measurements in the assessment of pain clinic results. *Acta Anaesthesiol. Scand.*, 1999, vol. 43, no. 9, pp. 893–896.
11. Indeks obshchego (khoroshego) samochuvstviia VOZ [WHO well-being index]. *Psychiatric Research Unit. WHO Collaborating Center for Mental Health, Frederiksborg General Hospital*, 1999. Available

at: https://www.psykiatri-regionh.dk/who-5/Documents/WHO5_Russian.pdf (accessed: 15.08.2017). (In Russian)

12. Diener E., Emmons R.A., Larsen R.J., Griffin S. The Satisfaction With Life Scale. *J. Pers. Assess.* Lawrence Erlbaum Associates, Inc. 1985, vol. 49, no. 1, pp. 71–75.

13. Ledovaia Ia. A., Bogolyubova O.N., Tikhonov R.V. Stress, blagopoluchie i Temnaia triada [Stress, Well-being and the Dark Triad]. *Psikhologicheskie issledovaniia*, 2015, vol. 8, no. 43, p. 5. Available at: <http://psystudy.ru/index.php/num/2015v8n43/1185-ledovaya43.html> (accessed: 01.09.2017). (In Russian)

14. Jones D.N., Paulhus D.L. Introducing the Short Dark Triad (SD3). *Assessment*, 2014, vol. 21, no. 1, pp. 28–41.

15. Egorova M.S., Sitnikova M.A., Parshikova O.V. Adaptatsiia Korotkogo oprosnika Temnoi triady [Adaptation of the Short Dark Triad]. *Psikhologicheskie issledovaniia*, 2015, vol. 8, no. 43, p. 1. Available at: <http://psystudy.ru/num/2015v8n43/1181-egorova43.html> (accessed: 01.09.2017). (In Russian)

16. Bandura A. *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice-Hall Inc, 1985. 544 p.

17. Moore C., Detert J.R., Klebe Treviño L., Baker V.L., Mayer D.M. Why employees do bad things: Moral disengagement and unethical organizational behavior. *Pers. Psychol.*, 2012, vol. 65, no. 1, pp. 1–48.

18. Ledovaia Ia.A., Tikhonov R.V., Bogoliubova O.N., Kazennaia E.V., Sorokina Iu.L. Otchuzhdenie moral'noi otvetstvennosti: psikhologicheskii konstrukt i metody ego izmereniia [Moral Disengagement: the Psychological Construct and its Measurement]. *Vestnik SPbSU. Series 16. Psychology. Education*, 2016, issue 4, pp. 23–39. (In Russian)

19. Prins A., Ouimette P., Kimerling R., Camerond R.P., Hugelshofer D.S., Shaw-Hegwer J., Sheikh J.I. The primary care PTSD screen (PC-PTSD): development and operating characteristics. *Prim. Care Psychiatry*, 2004, vol. 9, no. 1, pp. 9–14.

20. Bogolyubova O., Tikhonov R., Ivanov V., Panicheva P., Ledovaya Ya. Violence Exposure, Posttraumatic Stress, and Subjective Well-Being in a Sample of Russian Adults. *J. Interpers. Violence*, 2017, pp. 1–16. Available at: <https://doi.org/10.1177/0886260517698279> (accessed: 01.10.2017).

21. Bogolyubova O., Panicheva P., Tikhonov R., Ivanov V., Ledovaya Ya. Dark Personalities on Facebook: Harmful Online Behaviors and Language. *Comput. Human Behav. Elsevier Ltd*, 2018, vol. 78, pp. 151–159. Available at: <https://doi.org/10.1016/j.chb.2017.09.032> (accessed: 01.10.2017).

22. Panicheva P., Ledovaya Ya., Bogolyubova O. Lexical, Morphological and Semantic Correlates of the Dark Triad Personality Traits in Russian Facebook Texts. *Proc. Artif. Intell. Nat. Lang. Conf. (AINL 2016)*, 2016, pp. 1–8. Available at: <http://ieeexplore.ieee.org/document/7891860/> (accessed: 01.10.2017).

23. Panicheva P., Mirzagitova A., Ledovaya Y. Semantic Feature Aggregation for Gender Identification in Russian Facebook. *AINL 2017: Artificial Intelligence and Natural Language*. Eds A. Filchenkov, L. Pivovarova, J. Žižka. Communications in Computer and Information Science, vol. 789. Springer, Cham. 2018, pp. 3–15. URL: https://link.springer.com/chapter/10.1007/978-3-319-71746-3_1 (accessed: 01.12.2017).

24. Moskvichev A., Dubova M., Menshov S., Filchenkov A. Using Linguistic Activity in Social Networks to Predict and Interpret Dark Psychological Traits. *AINL 2017: Artificial Intelligence and Natural Language*. Eds A. Filchenkov, L. Pivovarova, J. Žižka. Communications in Computer and Information Science, vol. 789. Springer, Cham. 2018, pp. 16–26. URL: https://link.springer.com/chapter/10.1007/978-3-319-71746-3_2 (accessed: 01.12.2017).

25. Crowne D.P., Marlowe D. A new scale of social desirability independent of psychopathology. *J. Consult. Psychol.*, 1960, vol. 24, no. 4, pp. 349–354.

26. Kolichestvo aktivnykh pol'zovatelei Facebook v Rossii sostavliaet 14,4 mln chelovek [The number of active Facebook users in Russia is 14.4 million people]. *Rambler News Service*, 2016. Available at: <https://rns.online/internet/Kolichestvo-aktivnih-polzovatelei-Facebook-v-Rossii-sostavlyayet-144-mln-chelovek-2016-05-13/> (accessed: 01.09.2017). (In Russian)

27. Guðmundsdóttir H., Ólason D., Guðmundsdóttir D., Sigurðsson J.F. A psychometric evaluation of the Icelandic version of the WHO-5. *Scand. J. Psychol.*, 2014, vol. 55, no. 6, pp. 567–572.

For citation: Ledovaya Ya.A., Tikhonov R.V., Ivanov V.Yu., Yaminov B.R., Bogolyubova O.N. Organisational and methodological issues of data collection in an internet based study of Facebook users from Russia and USA. *Vestnik SPbSU. Psychology and Education*, 2017, vol. 7, issue 4, pp. 308–327. <https://doi.org/10.21638/11701/spbu16.2017.402>

Статья поступила в редакцию 12 октября 2017 г.

Статья принята к публикации 26 октября 2017 г.

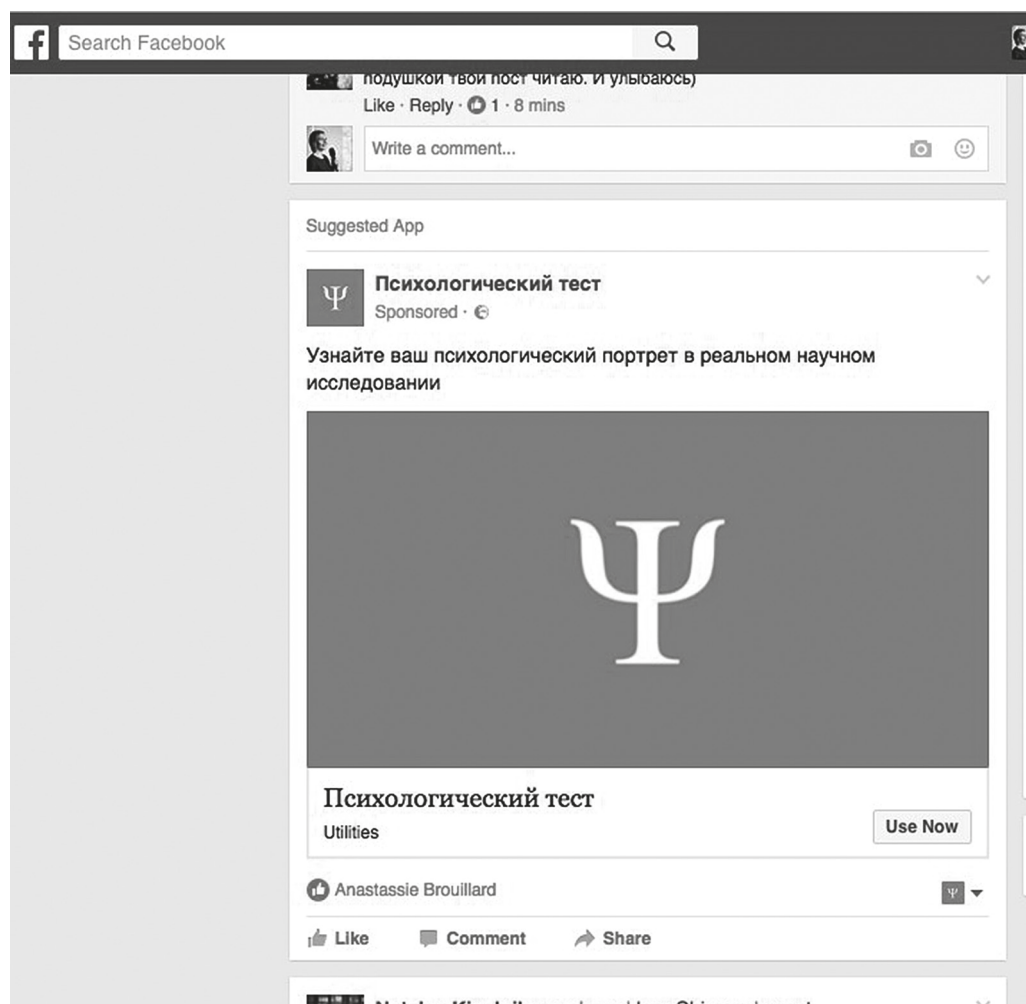


Рис. 1. Пример отображения рекламного объявления в ленте новостей социальной сети «Фейсбук»

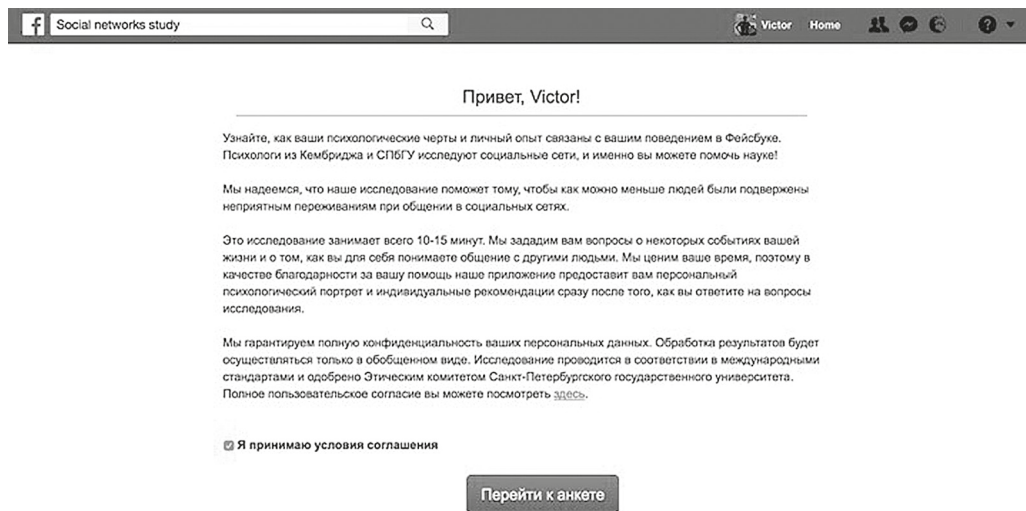


Рис. 2. Пример начальной страницы приложения с краткой формой информированного согласия

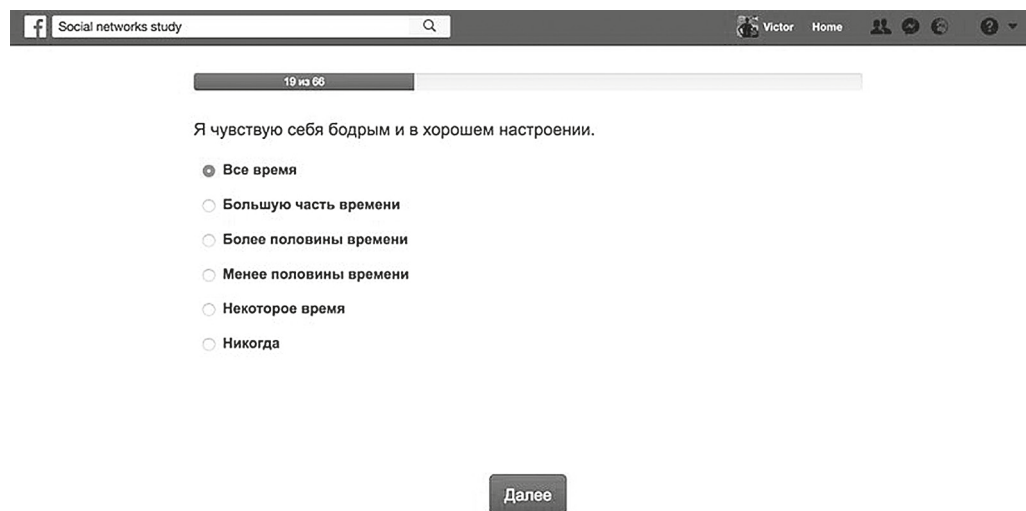


Рис. 3. Пример отображения вопросов и вариантов ответа для участников

Спасибо за ответы!

У вас 0 постов за последний месяц. Это меньше, чем у среднестатистического пользователя Facebook вашего возраста. Согласно исследованиям пользователи в среднем размещают 7 постов в месяц.

Ваши ответы близки с ответами других жителей вашего города. В сравнении с жителями вашего города, жители Москвы демонстрируют более высокий уровень стресса и чаще становятся жертвами троллинга.

Как правило, вы чувствуете себя нормально, но время от времени вам может казаться, что вам не хватает отдыха, вы можете чувствовать тревогу, напряжение, подавленность или апатию.

По своим личностным чертам вы больше всего похожи на **Афродиту** — древнегреческую богиню красоты и вечной весны, безжалостную к тем, кто отвергает любовь, и ставшую символом обольщения: в нее влюблялись практически все, кто встречался на ее пути.



Вы высоко оцениваете себя, свою внешность и достижения. Вы любите находиться в центре внимания и считаете, что ни одно мероприятие не будет стоящим без вас. Для вас важно постоянно подтверждать собственную значимость перед окружающими, вы очень не любите критику в свой адрес и можно сказать, что не очень трудолюбивы. Тем не менее, вы считаете, что заслуживаете особого отношения и уважения к себе от окружающих.

Поделиться

По всем вопросам вы можете связаться с нами - командой исследователей СПбГУ - по адресу: wellbeing.and.networks@gmail.com.

Рис. 4. Обратная связь для женщины-участницы, набравшей по шкале «нарциссизм» больше баллов, чем по двум другим шкалам «Темной триады»